Appendix 1

CALCULATION OF CONFIDENCE INTERVALS

1. **INTRODUCTION**

The use of Noise-Power-Distance (NPD) curves requires that confidence intervals be determined using a more general formulation than is used for a cluster of data points. For this more general case confidence intervals may need to be calculated about a regression line for:

(a) flight test data,

(b) a combination of flight test and static test data,

(c) analytical results,

or a combination thereof.

The latter two are of particular significance for noise certifications of an aircraft model range and require special care when pooling the different sources of sampling variability.

Sections 2 to 5 provide an insight into the theory of confidence interval evaluation. The application of this theory and some worked examples are presented in Section 6. A suggested bibliography is given in Section 7 for those wishing to gain a greater understanding.

2. **CONFIDENCE INTERVAL FOR THE MEAN OF FLIGHT TEST DATA**

2.1 **Confidence interval for the sample estimate of the mean of clustered measurements**

If *n* measurements of effective perceived noise levels $y_1, y_2, \ldots, y_n$ are obtained under approximately the same conditions <u>and</u> it can be assumed that they constitute a random sample from a normal population with true population mean, μ, and true standard deviation, σ, then the following statistics can be derived:-

$$\overline{y} = \text{estimate of the mean} = \frac{1}{n}\left\{\sum_{i=1}^{i=n} y(i)\right\},$$

$$s = \text{estimate of the standard deviation}$$

$$= \sqrt{\frac{\sum_{i=1}^{i=n}(y_i - \overline{y})^2}{n-1}}.$$

From these and the Student's t-distribution, the confidence interval, CI, for the estimate of the mean, $\overline{y}$, can be determined, as:

$$CI = \overline{y} \pm t_{\left(1-\frac{a}{2},z\right)} \frac{s}{\sqrt{n}}$$

where $t_{\left(1-\frac{a}{2},z\right)}$ denotes the $\left(1 - \dfrac{a}{2}\right)$ percentile of the single-sided Student's t-test with z degrees of freedom (for a clustered data set $z = n-1$) and where α is defined such that $100(1-a)$ percent is the desired confidence level for the confidence interval. That is it denotes the probability with which the interval will contain the unknown mean, μ. For noise certification purposes 90% confidence intervals are

generally desired and, thus $t_{.95,\mathbf{z}}$ is used. See Table 1-2 situated at the end of this appendix for a listing of values of $t_{.95,\mathbf{z}}$ for different values of ζ.

2.2 Confidence interval for mean Line obtained by regression

If n measurements of effective perceived noise levels $y_1, y_2, \ldots, y_n$ are obtained under significantly varying values of engine-related parameter $x_1, x_2, \ldots, x_n$ respectively, then a polynomial can be fitted to the data by the method of least squares. The following polynomial regression model for the mean effective perceived noise level, $\mathbf{m}$, is assumed to apply:

$$\mathbf{m} = B_0 + B_1 x + B_2 x^2 + \ldots + B_k x^k$$

and the <u>estimate of the mean</u> line through the data of the effective perceived noise level is given by:

$$y = b_0 + b_1 x + b_2 x^2 + \ldots + b_k x^k$$

Each regression coefficient B_i is estimated by b_i from the sample data using the method of least squares in a process summarised below.

Every observation (x_i, y_i) satisfies the equations

$$\begin{aligned} y_i &= B_0 + B_1 x_i + B_2 x_i^2 + \ldots + B_k x_i^k + \mathbf{e}_i \\ &= b_0 + b_1 x_i + b_2 x_i^2 + \ldots + b_k x_i^k + e_i \end{aligned}$$

where $\mathbf{e}_i$ and e_i are the random error and residual associated with the effective perceived noise level. The random error, $\mathbf{e}_i$, is assumed to be a random sample from a normal population with mean zero and standard deviation $\mathbf{s}$. The residual, e_i, is the difference between the measured value and the estimate of the value using the estimates of the regression coefficients and x_i. Its root mean square value, s, is the sample estimate for $\mathbf{s}$. These equations are often referred to as the normal equations.

The n data points of measurements (x_i, y_i) are processed as follows:

Each elemental vector, $\underline{x}_i$, and its transpose $\underline{x}'_i$, are formed such that

$$\underline{x}_i = \begin{pmatrix} 1 & x_i & x_i^2 & . & . & x_i^k \end{pmatrix} \text{ , a row vector,}$$

and

$$\underline{x}'_i = \begin{pmatrix} 1 \\ x_i \\ x_i^2 \\ . \\ . \\ x_i^k \end{pmatrix} \text{ , a column vector.}$$

A matrix $\underline{X}$ is formed from all the elemental vectors $\underline{x}_i$ for $i = 1, \ldots, n$. $\underline{X}'$ is the transpose of $\underline{X}$.

We define a matrix $\underline{A}$ such that $\underline{A} = \underline{X}' \underline{X}$ and a matrix $\underline{A}^{-1}$ to be the inverse of $\underline{A}$.

Also $\quad \underline{y} = \begin{pmatrix} y_1 & y_2 & . & . & . & y_n \end{pmatrix}$

and $\quad \underline{b} = \begin{pmatrix} b_0 & b_1 & . & . & . & b_k \end{pmatrix}$

with $\underline{b}$ determined as the solution of the normal equations:

$$\underline{y} = \underline{X}\underline{b}$$

and $\quad \underline{X'}\underline{y} = \underline{X'}\underline{X}\underline{b} = \underline{A}\underline{b} \quad,$

to give $\quad \underline{b} = \underline{A}^{-1}\underline{X'}\underline{y} \quad.$

The 90% confidence interval, CI_{90}, for the mean value of the effective perceived noise level estimated with the associated value of the engine-related parameter, x_0, is then defined as

$$CI_{90} = \overline{y}(x_0) \pm t_{.95,z} \ s \ v(x_0) \ , \text{ where } v(x_0) = \sqrt{\underline{x}_0 \underline{A}^{-1} \underline{x'}_0} \quad .$$

Thus $\quad CI_{90} = \overline{y}(x_0) \pm t_{.95,z} \ s \ \sqrt{\underline{x}_0 \underline{A}^{-1} \underline{x'}_0}$

where $\quad \underline{x}_0 = \begin{pmatrix} 1 & x_0 & x_0^2 & . & . & x_0^k \end{pmatrix},$

$\underline{x'}_0$ is the transpose of $\underline{x}_0$,

$\overline{y}(x_0)$ is the estimate of the mean value of the effective perceived noise level at the associated value of the engine related parameter,

$t_{.95,z}$ is obtained for $\mathbf{z}$ degrees of freedom. For the general case of a multiple regression analysis involving K independent variables (i.e. $K+1$ coefficients) $\mathbf{z}$ is defined as $\mathbf{z} = n - K - 1$ (for the specific case of a polynomial regression analysis, for which k is the order of curve fit, we have k variables independent of the dependent variable, and so $\mathbf{z} = n - k - 1$),

and $\quad s = \sqrt{\dfrac{\sum\limits_{i=1}^{i=n}\left(y_i - \overline{y}(x_i)\right)^2}{n - K - 1}}$, the estimate of $\mathbf{s}$, the true standard deviation.

3. CONFIDENCE INTERVAL FOR STATIC TEST DERIVED NPD CURVES

When static test data is used in family certifications, NPD curves are formed by the linear combination of baseline flight regressions, baseline projected static regressions, and derivative projected static regressions in the form:

$$EPNL_{DF} = EPNL_{BF} - EPNL_{BS} + EPNL_{DS}$$

or using the notation adopted above:

$$\overline{y}_{DF}(x_0) = \overline{y}_{BF}(x_0) - \overline{y}_{BS}(x_0) + \overline{y}_{DS}(x_0)$$

where subscript *DF* denotes derivative flight, *BF* denotes baseline flight, *BS* denotes baseline static, and *DS* denotes derivative static.

Confidence intervals for the derivative flight NPD curves are obtained by pooling the three data sets (each with their own polynomial regression). The confidence interval for the mean derived effective perceived noise level at engine-related parameter x_0, i.e., for $\mathbf{m}_{DF}(x_0)$, is given by:-

$$CI_{90}(x_0) = \overline{y}_{DF}(x_0) \pm t' v_{DF}(x_0)$$

where $v_{DF}(x_0) = \sqrt{\left(s_{BF}v_{BF}(x_0)\right)^2 + \left(s_{BS}v_{BS}(x_0)\right)^2 + \left(s_{DS}v_{DS}(x_0)\right)^2}$

with s_{BF}, s_{BS}, s_{DS}, $v_{BF}(x_0)$, $v_{BS}(x_0)$ and $v_{DS}(x_0)$ computed as explained in Section 2.2 for the respective data sets indicated by the subscripts *BF*, *BS*, and *DS*, and

$$t' = \frac{\left(s_{BF}v_{BF}(x_0)\right)^2 t_{BF} + \left(s_{BS}v_{BS}(x_0)\right)^2 t_{BS} + \left(s_{DS}v_{DS}(x_0)\right)^2 t_{DS}}{\left(s_{BF}v_{BF}(x_0)\right)^2 + \left(s_{BS}v_{BS}(x_0)\right)^2 + \left(s_{DS}v_{DS}(x_0)\right)^2}$$

where t_{BF}, t_{BS} and t_{DS} are the $t_{.95,z}$ values each evaluated with the respective degrees of freedom z_{BF}, z_{BS} and z_{DS} as they arise in the corresponding regressions.

4. CONFIDENCE INTERVAL FOR ANALYTICALLY DERIVED NPD CURVES

Analysis may be used to determine the effect of changes in noise source components on certified levels. This is accomplished by analytically determining the effect of hardware change on the noise component it generates. The resultant delta is applied to the original configuration and new noise levels are computed. The changes may occur on the baseline configuration or on subsequent derivative configurations. The confidence intervals for this case are computed using the appropriate method from above. If $\overset{\backprime}{\Delta}$ represents the analytically determined change and if it is assumed that it may deviate from the true unknown Δ by some random amount d, i.e.

$$\overset{\backprime}{\Delta} = \Delta + d \quad,$$

where d is assumed to be normally distributed with mean zero and known variance t^2,

then the confidence interval for $m(x_0) + \Delta$ is given by

$$\left(\overline{y}(x_0) + \overset{\backprime}{\Delta}\right) \pm t'v'(x_0)$$

where $v'(x_0) = \sqrt{v(x_0)^2 + t^2}$ and t' is as above without change.

5. ADEQUACY OF THE MODEL

5.1 Choice of engine-related parameter

Every effort should be made to determine the most appropriate engine-related parameter *x*, which may be a combination of various simpler parameters.

5.2 Choice of regression model

It is not recommended in any case that polynomials of greater complexity than a simple quadratic be used for certification purposes, unless there is a clear basis for such a model.

Standard texts on multiple regression should be consulted and the data available should be examined to show the adequacy of the model chosen.

6. **WORKED EXAMPLE OF THE DETERMINATION OF 90% CONFIDENCE INTERVALS FROM THE POOLING OF THREE DATA SETS**

This section presents an example of the derivation of the 90% confidence intervals arising from the pooling of three data sets. Worked examples and guidance material are presented for the calculation of confidence intervals for a clustered data set and for first order (ie. straight line) and second order (ie. quadratic) regression curves. In addition it is shown how the confidence interval shall be established for the pooling together of several data sets.

Consider the theoretical evaluation of the certification noise levels for an aircraft retro-fitted with silenced engines. The approach noise level for the datum aircraft was derived from a clustered data set of noise levels measured at nominally reference conditions, to which were added source noise corrections derived from a quadratic least squares curve fit through a series of data points made at different engine thrusts. In order to evaluate the noise levels for the aircraft fitted with acoustically treated engines a further source noise curve (assumed to be a straight least squares regression line) was established from a series of measurements of the silenced aircraft. Each of the three data bases is assumed to be made up of data unique to each base.

The clustered data set consists of six EPNL levels for the nominal datum hardwall condition. These levels have been derived from measurements which have been fully corrected to the hardwall approach reference condition.

The two curves which determine the acoustic changes are the regression curves (in the example given both a quadratic and straight line least squares curve fit) for the plots of EPNL against normalised thrust for the hardwall and silenced conditions. These are presented in figure below. The dotted lines plotted about each line represent the boundaries of 90% confidence.
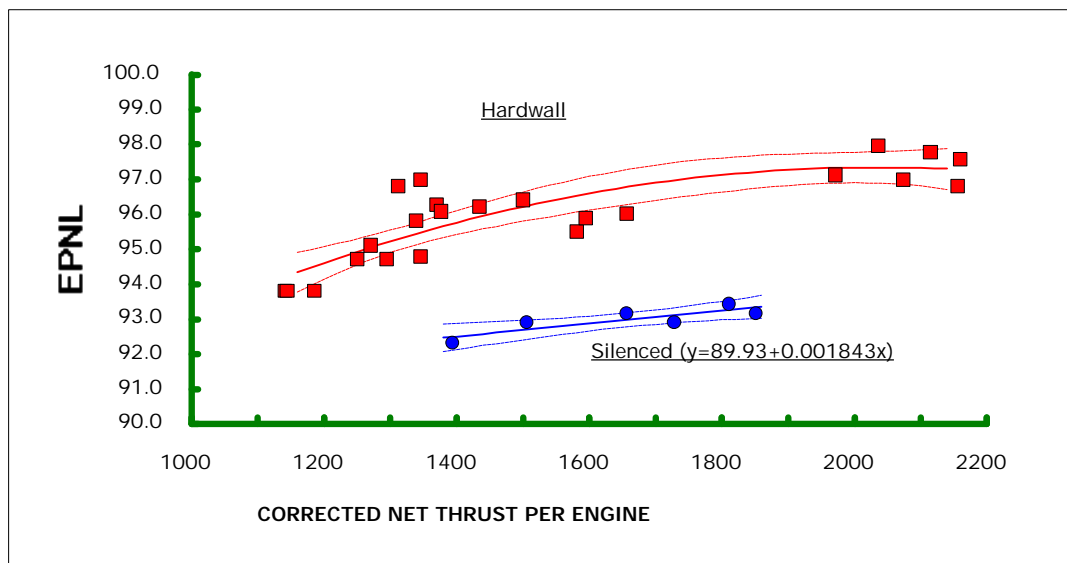

Figure 1.1

Each of the two curves is made up of from the full set of data points obtained for each condition during a series of back to back tests. The least squares fits therefore have associated with them all the uncertainties contained within each data set. It is considered that the number of data points in each of the three sets is large enough to constitute a statistical sample.

6.1 **Confidence interval for a clustered data set**

The confidence interval of the clustered data set is defined as follows:

Let $EPNL_i$ be the individual values of EPNL
n = number of data points
t = Student's t-distribution for (n-1) degrees of freedom (the number of degrees of

freedom associated with a clustered data set).

Then the Confidence Interval $CI = \overline{EPNL} \pm t \dfrac{s}{\sqrt{n}}$

where s, the estimate of the standard deviation, is defined as

$$s = \sqrt{\frac{\sum\limits_{i=1}^{i=n}\left(EPNL_i - \overline{EPNL}\right)^2}{n-1}}$$

and $\quad \overline{EPNL} = \dfrac{\sum\limits_{i=1}^{i=n} EPNL_i}{n}$.

Let us suppose that our clustered set of EPNL values consists of the following:

Run Number	EPNL
1	95.8
2	94.8
3	95.7
4	95.1
5	95.6
6	95.3

Then   number of data points $(n) = 6$,

degrees of freedom $(n-1) = 5$,

Student's t-distribution for 5 degrees of freedom $= 2.015$ (See Table 1-2),

$$\overline{EPNL} = \frac{\sum\limits_{i=1}^{i=n} EPNL_i}{n} = 95.38,$$

$$s = \sqrt{\frac{\sum\limits_{i=1}^{i=n}\left(EPNL_i - \overline{EPNL}\right)^2}{n-1}} = 0.3869 ,$$

and Confidence Interval

$$CI = \overline{EPNL} \pm t \frac{s}{\sqrt{n}} = 95.38 \pm 2.015 \frac{0.3869}{\sqrt{6}} = 95.38 \pm \underline{\underline{0.3183}}$$

6.2 **Confidence interval for a first order regression curve**

Let us suppose that the regression curve for one of the source noise data sets (for the silenced case) can best be represented by a least squares straight line fit ie. a first order polynomial.

The equation for this regression line is of the general form:

$$Y = a + bX$$

where Y represents the dependent variable $EPNL$,

and X represents the independent variable normalised thrust F_N/d ,(in this case).

Although for higher order polynomial least squares curves a regression line's coefficients (ie. the solutions to the "normal equations") are best established through computer matrix solutions, the two coefficients for a straight line fit, a and b, can be determined from the following two simple formulae for the measured values of X and Y, X_i and Y_i:

$$b = \frac{Covariance}{Variance} = \frac{S_{xy}^2}{S_x^2}$$

where $\quad S_{xy}^2 = \dfrac{\sum\limits_{i=1}^{i=n} X_i Y_i}{n} - \dfrac{\sum\limits_{i=1}^{i=n} X_i \sum\limits_{i=1}^{i=n} Y_i}{n^2}$

and $\quad S_x^2 = \dfrac{\sum\limits_{i=1}^{i=n} X_i^2}{n} - \left(\dfrac{\sum\limits_{i=1}^{i=n} X_i}{n}\right)^2$,

$$a = \frac{\sum\limits_{i=1}^{i=n} Y_i - b \sum\limits_{i=1}^{i=n} X_i}{n}$$

The 90% confidence interval about this regression line for $X = x_0$ is then defined by:

$$CI_{90} = \overline{Y} \pm ts\sqrt{\underline{x_0}\,\underline{A}^{-1}\,\underline{x_0}'}$$

where t = Student's t-distribution for 90% confidence corresponding to $(n - k - 1)$ degrees of freedom (where k is the order of the polynomial regression line and n is the number of data points),

$\underline{x_0} = \begin{pmatrix} 1 & x_0 \end{pmatrix}$ and $\underline{x_0}' = \begin{pmatrix} 1 \\ x_0 \end{pmatrix}$,

$\underline{A}^{-1}$ is the inverse of $\underline{A}$ where $\underline{A} = \underline{X}'\underline{X}$,

with $\underline{X}$ and $\underline{X}'$ defined as in paragraph 2.2 from the elemental vectors formed from the measured values of independent variable X_i ,

$$s = \sqrt{\frac{\sum\limits_{i=1}^{i=n} (\Delta Y)_i^2}{n - k - 1}}$$

where $(\Delta Y)_i$ = the difference between the measured value of Y_i at its associated value of X_i ,and the value of Y derived from the least squares fit straight line for $X = X_i$,and n and k are defined as above.

Let us suppose that our data set consists of the following set of six EPNL values together with their associated values of engine related parameter (Note that it would be usual to have more than six data points making up a source noise curve but in order to limit the size of the matrices in this example their number has been restricted):

Run Number	F_N/d	EPNL
1	1395	92.3
2	1505	92.9
3	1655	93.2
4	1730	92.9
5	1810	93.4
6	1850	93.2

Table 1-1

By plotting this data (See Figure 1.1) it can be seen by examination that a linear relationship between EPNL (the dependent variable Y) and F_N/d (the independent variable X) is suggested with the following general form:

$$Y = a + bX$$

The coefficients a and b of the linear equation are defined as above and may be calculated as follows:

X	Y	XY	X^2
1395	92.3	128759	1946025
1505	92.9	139815	2265025
1655	93.2	154246	2739025
1730	92.9	160717	2992900
1810	93.4	169054	3276100
1850	93.2	172420	3422500
$\sum X$	$\sum Y$	$\sum XY$	$\sum X^2$
9945	557.9	925010	16641575

$$b = \frac{Covariance}{Variance} = \frac{S_{xy}{}^2}{S_x{}^2} \quad \text{where}$$

$$S_{xy}{}^2 = \frac{\sum_{i=1}^{i=n} X_i Y_i}{n} - \frac{\sum_{i=1}^{i=n} X_i \sum_{i=1}^{i=n} Y_i}{n^2} = \frac{925010}{6} - \frac{(9945)(557.9)}{36} = 48.46$$

and $\quad S_x{}^2 = \dfrac{\sum_{i=1}^{i=n} X_i{}^2}{n} - \left(\dfrac{\sum_{i=1}^{i=n} X_i}{n} \right)^2 = \dfrac{16641575}{6} - \left(\dfrac{9945}{6} \right)^2 = 26289.6$

to give $\quad b = \dfrac{48.46}{26289.6} = \underline{\underline{0.001843}}$

and $\quad a = \dfrac{\sum_{i=1}^{i=n} Y_i - b \sum_{i=1}^{i=n} X_i}{n} = \dfrac{557.9 - (0.001843)(9945)}{6} = \underline{\underline{89.93}}$

The 90% confidence interval about this regression line which is defined as:

$$CI_{90} = \overline{Y} \pm ts\sqrt{\underline{x_0}\,\underline{A}^{-1}\,\underline{x_0}'}$$

is calculated as follows.

From the single set of measured independent variables tabulated in Table 1-1 let us form the matrix, $\underline{X}$, from the elemental row vectors such that:

$$\underline{X} = \begin{pmatrix} 1 & 1395 \\ 1 & 1505 \\ 1 & 1655 \\ 1 & 1730 \\ 1 & 1810 \\ 1 & 1850 \end{pmatrix},$$

and $\underline{X}'$, the transpose of $\underline{X}$, where

$$\underline{X}' = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1395 & 1505 & 1655 & 1730 & 1810 & 1850 \end{pmatrix}.$$

We now form the matrix $\underline{A}$, defined such that $\underline{A} = \underline{X}'\,\underline{X}$, and so

$$\underline{A} = \begin{pmatrix} 6 & 9945 \\ 9945 & 16641575 \end{pmatrix} \text{ and its inverse } \underline{A}^{-1} \text{ such that}$$

$$\underline{A}^{-1} = \begin{pmatrix} 17.5836 & -0.01051 \\ -0.01051 & 6.3396\text{E}-6 \end{pmatrix}.$$

NB. The manipulation of matrices, their multiplication and inversion, are best performed by computers via standard routines. Such routines are possible using standard functions contained within many commonly used spreadsheets.

Suppose for example we now wish to find the 90% confidence interval about the regression line for a value of F_N/d (i.e. x_0) of 1600. We form the row vector $\underline{x_0}$ such that:

$$\underline{x_0} = \begin{pmatrix} 1 & 1600 \end{pmatrix} \text{ and its transpose , a column vector } \underline{x_0}' = \begin{pmatrix} 1 \\ 1600 \end{pmatrix}.$$

From our calculation of $\underline{A}^{-1}$ we have:

$$\underline{x_0}\,\underline{A}^{-1} = \begin{pmatrix} 1 & 1600 \end{pmatrix} \begin{pmatrix} 17.5836 & -0.01051 \\ -0.01051 & 6.3396\text{E}-6 \end{pmatrix}$$

$$= \begin{pmatrix} 0.7709 & -3.6453\text{E}-4 \end{pmatrix}$$

and so $\quad \underline{x_0}\,\underline{A}^{-1}\,\underline{x_0}' = \begin{pmatrix} 0.7709 & -3.6453\text{E}-4 \end{pmatrix} \begin{pmatrix} 1 \\ 1600 \end{pmatrix}$

$$= 0.1876 \ .$$

Our equation for confidence interval also requires that we evaluate the value of standard deviation for the measured data set. From Table 1-1 and our regression equation for the least squares best fit straight line (from which we calculate the predicted value of EPNL at each of the 6 measured values of F_N/d) we proceed as follows:

Run Number	F_N/d	EPNL (Measured)	EPNL (Predicted)	$(\Delta EPNL)^2$
1	1395	92.3	92.50	0.03979
2	1505	92.9	92.70	0.03911
3	1655	93.2	92.98	0.04896
4	1730	92.9	93.12	0.04708
5	1810	93.4	93.26	0.01838
6	1850	93.2	93.34	0.01909

$$s = \sqrt{\dfrac{\sum\limits_{i=1}^{i=n}(\Delta y)_i^2}{n-k-1}} = \sqrt{\dfrac{0.21241}{6-1-1}} = 0.2304 \ \ \text{for} \ \ n=6 \ \ \text{and} \ \ k=1.$$

and so taking the value of Student's t from Table 1-2 for $(n-k-1)$ degrees of freedom (i.e. 4) to be 2.132, we have the confidence interval about the regression line at $F_N/d = 1600$ defined as follows:

$$CI_{90} = \overline{\text{EPNL}} \pm \text{t}\ s\sqrt{\underline{x_0}\,\underline{A}^{-1}\,\underline{x_0}'} = 92.98 \pm (2.132)(0.2304)\sqrt{0.1876} = 92.98 \pm \underline{\underline{0.2128}}$$

In order to establish the lines of 90% confidence intervals about a regression line the values of CI_{90} for a range of values of independent variable(s) should be calculated, through which a line can be drawn. These lines are shown as the dotted lines on Figure 1.1

6.3 **Confidence interval for a second order regression curve**

The confidence intervals about a second order regression curve are derived in a similar manner to those for a straight line detailed in Section 6.1. It is not felt that a detailed example of their calculation would be appropriate. However the following points should be borne in mind.

The coefficients of the least squares regression quadratic line are best determined via computer matrix solutions. Regression analysis functions are a common feature of many proprietary software packages.

The matrices $\underline{x_0}$, $\underline{x_0}'$, $\underline{X}$ and $\underline{X}'$ formed during the computation of the confidence interval according to the formula:

$$CI_{90} = \overline{Y} \pm t\ s\sqrt{\underline{x_0}\,\underline{A}^{-1}\,\underline{x_0}'}$$

are formed from 1 x 3 and 3 x 1 row and column vectors respectively, made up from the values of independent variable X according to the following general form:

$$\underline{x} = \begin{pmatrix} 1 & x & x^2 \end{pmatrix} \text{ and } \underline{x}' = \begin{pmatrix} 1 \\ x \\ x^2 \end{pmatrix}.$$

The number of degrees of freedom associated with a multiple regression analysis involving K variables independent of the dependent variable (ie. with $(K+1)$ coefficients, including the constant term) is defined as $(n-K-1)$. For a second order regression curve we have two independent variables and so the number of degrees of freedom is $(n-3)$.

6.4 **Confidence interval for the pooled data set**

The confidence interval associated with the pooling of three data sets is defined as follows:

$$CI = \overline{Y} \pm T' \sqrt{\sum_{i=1}^{i=3} Z_i^2}$$

where $\quad Z_i = \dfrac{CI_i}{t_i}$, with CI_i = confidence interval for the i'th data set and t_i = value of Student's

t for the i'th data set,

and $\quad T_i = \dfrac{\sum\limits_{i=1}^{i=3} Z_i^2 t_i}{\sum\limits_{i=1}^{i=3} Z_i^2}$.

The different stages in the calculation of the confidence interval at our reference thrust of $F_N/d = 1600$ for the pooling of our three data sets is summarised below:

Description	Function	Datum	Hardwall	Silenced
Reference Thrust	F_N/d		1600	1600
90% Confidence Interval about the mean	CI_{90}	0.3183	0.4817	0.2128
Number of data points	n	6	23	6
Degree of curve fit	k	0	2	1
Number of independent variables	K	0	2	1
Number of degrees of freedom	$n - K - 1$	5	20	4
Student's t	t	2.015	1.725	2.132
Z	CI_{90}/t	0.1580	0.2792	0.09981
Z^2	$\left(CI_{90}/t\right)^2$	2.4953E-2	7.7979E-2	9.9625E-3
$Z^2 t$	$\left(CI_{90}/t\right)^2 t$	5.0280E-2	0.1345	2.1240E-2
$\sum Z^2$		0.1129		
$\sum \left(Z^2 t\right)$		0.2060		
T	$\sum \left(Z^2 t\right) / \sum Z$	1.8248		
$\sqrt{\sum Z^2}$		0.3360		
CI	$T\sqrt{\sum Z^2}$	0.6131		

7. **BIBLIOGRAPHY**

1. Hafner, New York, 1971. Kendal, M.G. and Stuart, A., The Advanced Theory of Statistics, Volumes 1 and 2, 3

2. Yule, G.U. and Kendall, M.G., An Introduction to the Theory of Statistics, 14th ed., Griffin, New York, 1950.

3. Walpole, R.E. and Myers R.H., Probability and Statistics for Engineers and Scientists. MacMillan, New York, 1972.

4. Cochran, W.G., Approximate Significance Levels of the Behrens-Fisher Test, Biometrics, 10 191-195, 1964.

5. Rose, D.M. and Scholz, F.W., Statistical Analysis of Cumulative Shipper-Receiver Data, NUREG/CR-2819, Division of Facility Operations, Office of Nuclear Regulatory Research, U.S. Nuclear Regulatory Commission, Washington, D.C. 20555, NRC FIN B1076, 1983.

6. Snedecor, G.W. and Cochran, W.G., Statisitical Methods, 6th ed., The Iowa State University Press, Arnes, Iowa, 1968.

7. Wonnacott, T.H. and Wonnacott, R.J., Introductory Statistics, 5th ed., John Wiley & Sons, 1990

Degrees of Freedom (ζ)	$t_{.95,z}$
1	6.314
2	2.920
3	2.353
4	2.132
5	2.015
6	1.943
7	1.895
8	1.860
9	1.833
10	1.812
12	1.782
14	1.761
16	1.746
18	1.734
20	1.725
24	1.711
30	1.697
60	1.671
>60	1.645

Values in the Student's t-distribution to give a probability of 0.95 that the population mean value, m, is such that:

$$m \leq \bar{y} + t_{.95,z} \frac{s}{\sqrt{n}}$$, and thus a probability of 90% that

$$\bar{y} - t_{.95,z} \frac{s}{\sqrt{n}} \leq m \leq \bar{y} + t_{.95,z} \frac{s}{\sqrt{n}} \quad .$$

Student's t-DISTRIBUTION (FOR 90% CONFIDENCE INTERVAL)
FOR VARIOUS DEGREES OF FREEDOM

TABLE 1-2